Fig. 4. The impact of risk of anemia and RBV dose on treatment outcome after a complete early virological response. Patients with complete early virological responses were divided into subgroups according to their adherence to RBV: ≤40%, 41–60%, 61–80%, and >80%. For each subgroup, the proportion of patients with a high risk and a low risk of anemia is shown in the upper panel by pie charts, and the rates of sustained virological responses, stratified by high risk and low risk of anemia, are shown in the lower panel by bar graphs. The black and white bars or charts represent patients with high and low risks of anemia, respectively.

high risk of anemia to extend the duration of treatment, even those patients with a complete early virological response, to obtain >80% adherence to RBV.

In conclusion, the combination of the *ITPA* genotype, baseline Hb concentration, and baseline CLcr could be used as a pre-treatment predictor of anemia. The risk of anemia thus identified is associated with adherence to RBV and impacts on the treatment outcome of patients who achieve a complete early virological response. This is in contrast to the major role of the *IL28B* genotype in the prediction of sustained virological responses at baseline and among non-responders at weeks 4 and 12. Patients who achieve a complete early virological response generally have a high probability of a sustained virological response but those who have a high risk of anemia have a high rate of relapse because of reduced adherence to RBV. To improve the rate of sustained virological responses in these patients, it may be postulated that the treatment schedule may be personalized to obtain >80% adherence to RBV. Clearly, this postulate needs to be confirmed in a future study.

REFERENCES

Averbook BJ, Fu P, Rao JS, Mansour EG. 2002. A long-term analysis of 1018 patients with melanoma by classic Cox regression and tree-structured survival analysis at a major referral

center: Implications on the future of cancer staging. Surgery 132:589–602.

Azakami T, Hayes CN, Sezaki H, Kobayashi M, Akuta N, Suzuki F, Kumada H, Abe H, Miki D, Tsuge M, Imamura M, Kawakami Y, Takahashi S, Ochi H, Nakamura Y, Kamatani N, Chayama K. 2011. Common genetic polymorphism of ITPA gene affects ribavirin-induced anemia and effect of peg-interferon plus ribavirin therapy. J Med Virol 83:1048–1057.

Baquerizo A, Anselmo D, Shackleton C, Chen TW, Cao C, Weaver M, Gornbein J, Geevarghese S, Nissen N, Farmer D, Demetriou A, Busuttil RW. 2003. Phosphorus ans an early predictive factor in patients with acute liver failure. Transplantation 75:2007–2014.

Breiman LJH, Friedman RA, Olshen CJ, Stone CM. 1980. Classification and regression trees. Calif: Wadsworth.

Cockcroft DW, Gault MH. 1976. Prediction of creatinine clearance from serum creatinine. Nephron 16:31–41.

Davis GL, Wong JB, McHutchison JG, Manns MP, Harvey J, Albrecht J. 2003. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. Hepatology 38:645–652.

Fellay J, Thompson AJ, Ge D, Gumbs CE, Urban TJ, Shianna KV, Little LD, Qiu P, Bertelsen AH, Watson M, Warner A, Muir AJ, Brass C, Albrecht J, Sulkowski M, McHutchison JG, Goldstein DB. 2010. ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. Nature 464: 405–408.

Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales FL, Jr., Haussinger D, Diago M, Carosi G, Dhumeaux D, Craxi A, Lin A, Hoffman J, Yu J. 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 347:975–982.

Garzotto M, Park Y, Mongoue-Tchokote S, Bledsoe J, Peters L, Blank BH, Austin D, Beer TM, Mori M. 2005. Recursive partitioning for

risk stratification in men undergoing repeat prostate biopsies. Cancer 104:1911–1917.

Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 461:399–401.

Hiramatsu N, Kurashige N, Oze T, Takehara T, Tamura S, Kasahara A, Oshita M, Katayama K, Yoshihara H, Imai Y, Kato M, Kawata S, Tsubouchi H, Kumada H, Okanoue T, Kakumu S, Hayashi N. 2008. Early decline of hemoglobin can predict progression of hemolytic anemia during pegylated interferon and ribavirin combination therapy in patients with chronic hepatitis C. Hepatol Res 38:52–59.

Hiramatsu N, Oze T, Yakushijin T, Inoue Y, Igura T, Mochizuki K, Imanaka K, Kaneko A, Oshita M, Hagiwara H, Mita E, Nagase T, Ito T, Inui Y, Hijioka T, Katayama K, Tamura S, Yoshihara H, Imai Y, Kato M, Yoshida Y, Tatsumi T, Ohkawa K, Kiso S, Kanto T, Kasahara A, Takehara T, Hayashi N. 2009. Ribavirin dose reduction raises relapse rate dose-dependently in genotype 1 patients with hepatitis C responding to pegylated interferon alpha-2b plus ribavirin. J Viral Hepat 16: 586–594.

Hiramatsu N, Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Sugauchi F, Tamori A, Kakinuma S, Matsuura K, Izumi N. 2011. Pretreatment prediction of anemia progression by pegylated interferon alpha-2b plus ribavirin combination therapy in chronic hepatitis C infection: Decision-tree analysis. J Gastroenterol 46:1111–1119.

Kim RW. 2002. Global epidemiology and burden of hepatitis C. Microbes Infect 4:1219–1225.

Kurosaki M, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, Tamaki N, Hosokawa T, Ueda K, Tsuchiya K, Nakanishi H, Ikeda H, Itakura J, Takahashi Y, Asahina Y, Higaki M, Enomoto N, Izumi N. 2010. A predictive model of response to peginterferon ribavirin in chronic hepatitis C using classification and regression tree analysis. Hepatol Res 40:251–260.

Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, Sugauchi F, Tamori A, Nakagawa M, Izumi N. 2011a. Sequences in the interferon sensitivity-determining region and core region of hepatitis C virus impact pretreatment prediction of response to PEG-interferon plus ribavirin: Data mining analysis. J Med Virol 83:445–452.

Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, Sugauchi F, Yatsuhashi H, Izumi N. 2011b. Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis. J Gastroenterol 46:401–409.

Kurosaki M, Tanaka Y, Nishida N, Sakamoto N, Enomoto N, Honda M, Sugiyama M, Matsuura K, Sugauchi F, Asahina Y, Nakagawa M, Watanabe M, Sakamoto M, Maekawa S, Sakai A, Kaneko S, Ito K, Masaki N, Tokunaga K, Izumi N, Mizokami M. 2011c. Pretreatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in IL28B and viral factors. J Hepatol 54:439–448.

Kurosaki M, Tanaka Y, Tanaka K, Suzuki Y, Hoshioka Y, Tamaki N, Kato T, Yasui Y, Hosokawa T, Ueda K, Tsuchiya K, Kuzuya T, Nakanishi H, Itakura J, Takahashi Y, Asahina Y, Matsuura K, Sugauchi F, Enomoto N, Nishida N, Tokunaga K, Mizokami M, Izumi N. 2011d. Relationship between genetic polymorphisms of the inosine triphosphatase gene and anemia or outcome after treatment with pegylated-interferon and ribavirin. Antivir Ther 16:685–694.

Kurosaki M, Hiramatsu N, Sakamoto N, Suzuki Y, Iwasaki M, Tamori A, Matsuura K, Kakinuma S, Sugauchi F, Sakamoto N, Nakagawa M, Yatsuhashi H, Izumi N. 2012. Age and total ribavirin dose is an independent predictor of relapse among early virological responders to peg-interferon plus ribavirin therapy in chronic hepatitis C revealed by data mining analysis. Antivir Ther 17:35–43.

Kwo PY, Lawitz EJ, McCone J, Schiff ER, Vierling JM, Pound D, Davis MN, Galati JS, Gordon SC, Ravendhran N, Rossaro L, Anderson FH, Jacobson IM, Rubin R, Koury K, Pedicone LD, Brass CA, Chaudhri E, Albrecht JK. 2010. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): An open-label, randomised, multicentre phase 2 trial. Lancet 376:705–716.

LeBlanc M, Crowley J. 1995. A review of tree-based prognostic models. Cancer Treat Res 75:113–124.

Leiter U, Buettner PG, Eigentler TK, Garbe C. 2004. Prognostic factors of thin cutaneous melanoma: An analysis of the Central Malignant Melanoma Registry of the German Dermatological Society. J Clin Oncol 22:3660–3667.

Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, Goodman ZD, Koury K, Ling M, Albrecht JK. 2001. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: A randomised trial. Lancet 358:958–965.

McHutchison JG, Manns M, Patel K, Poynard T, Lindsay KL, Trepo C, Dienstag J, Lee WM, Mak C, Garaud JJ, Albrecht JK. 2002. Adherence to combination therapy enhances sustained response in genotype-1-infected patients with chronic hepatitis C. Gastroenterology 123:1061–1069.

McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, Kauffman R, McNair L, Alam J, Muir AJ. 2009. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. N Engl J Med 360:1827–1838.

Miyaki K, Takei I, Watanabe K, Nakashima H, Omae K. 2002. Novel statistical classification model of type 2 diabetes mellitus patients for tailor-made prevention using data mining algorithm. J Epidemiol 12:243–248.

Ochi H, Maekawa T, Abe H, Hayashida Y, Nakano R, Kubo M, Tsunoda T, Hayes CN, Kumada H, Nakamura Y, Chayama K. 2010. Inosine triphosphate pyrophosphatase polymorphism affects ribavirin-induced anemia and outcomes of therapy—A genome-wide study of Japanese hepatitis C virus patients. Gastroenterology 139:1190–1197.

Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, Bochud M, Battegay M, Bernasconi E, Borovicka J, Colombo S, Cerny A, Dufour JF, Furrer H, Gunthard HF, Heim M, Hirschel B, Malinverni R, Moradpour D, Mullhaupt B, Witteck A, Beckmann JS, Berg T, Bergmann S, Negro F, Telenti A, Bochud PY. 2010. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: A genome-wide association study. Gastroenterology 138:1338–1345.

Sakamoto N, Tanaka Y, Nakagawa M, Yatsuhashi H, Nishiguchi S, Enomoto N, Azuma S, Nishimura-Sakurai Y, Kakinuma S, Nishida N, Tokunaga K, Honda M, Ito K, Mizokami M, Watanabe M. 2010. ITPA gene variant protects against anemia induced by pegylated interferon-alpha and ribavirin therapy for Japanese patients with chronic hepatitis C. Hepatol Res 40: 1063–1071.

Seto WK, Tanaka Y, Liu K, Lai CL, Yuen MF. 2011. The effects of IL-28B and ITPA polymorphisms on treatment of hepatitis C virus genotype 6. Am J Gastroenterol 106:1007–1008.

Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E, Riordan S, Sheridan D, Smedile A, Fragomeli V, Muller T, Bahlo M, Stewart GJ, Booth DR, George J. 2009. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 41:1100–1104.

Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K, Masaki N, Sugauchi F, Izumi N, Tokunaga K, Mizokami M. 2009. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41:1105–1109.

Tanaka Y, Kurosaki M, Nishida N, Sugiyama M, Matsuura K, Sakamoto N, Enomoto N, Yatsuhashi H, Nishiguchi S, Hino K, Hige S, Itoh Y, Tanaka E, Mochida S, Honda M, Hiasa Y, Koike A, Sugauchi F, Kaneko S, Izumi N, Tokunaga K, Mizokami M. 2011. Genome-wide association study identified ITPA/DDRGK1 variants reflecting thrombocytopenia in pegylated interferon and ribavirin therapy for chronic hepatitis C. Hum Mol Genet 20:3507–3516.

Thompson AJ, Fellay J, Patel K, Tillmann HL, Naggie S, Ge D, Urban TJ, Shianna KV, Muir AJ, Fried MW, Afdhal NH, Goldstein DB, McHutchison JG. 2010a. Variants in the ITPA gene protect against ribavirin-induced hemolytic anemia and decrease the need for ribavirin dose reduction. Gastroenterology 139:1181–1189.

Thompson AJ, Muir AJ, Sulkowski MS, Ge D, Fellay J, Shianna KV, Urban T, Afdhal NH, Jacobson IM, Esteban R, Poordad F,

Lawitz EJ, McCone J, Shiffman ML, Galler GW, Lee WM, Reindollar R, King JW, Kwo PY, Ghalib RH, Freilich B, Nyberg LM, Zeuzem S, Poynard T, Vock DM, Pieper KS, Patel K, Tillmann HL, Noviello S, Koury K, Pedicone LD, Brass CA, Albrecht JK, Goldstein DB, McHutchison JG. 2010b. Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. Gastroenterology 139:120–129 e118.

Thompson AJ, Santoro R, Piazzolla V, Clark PJ, Naggie S, Tillmann HL, Patel K, Muir AJ, Shianna KV, Mottola L, Petruzzellis D, Romano M, Sogari F, Facciorusso D, Goldstein DB, McHutchison JG, Mangia A. 2011. Inosine triphosphatase genetic variants are protective against anemia during antiviral therapy for HCV2/3 but do not decrease dose reductions of RBV or increase SVR. Hepatology 53:389–395.

Valera VA, Walter BA, Yokoyama N, Koyama Y, Iiai T, Okamoto H, Hatakeyama K. 2007. Prognostic groups in colorectal carcinoma patients based on tumor cell proliferation and classification and regression tree (CART) survival analysis. Ann Surg Oncol 14:34–40.

Zlobec I, Steele R, Nigam N, Compton CC. 2005. A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. Clin Cancer Res 11:5440–5443.

PLoS one

Highly Parallel and Short-Acting Amplification with Locus-Specific Primers to Detect Single Nucleotide Polymorphisms by the DigiTag2 Assay

Nao Nishida[1,2]*, Yoriko Mawatari[1,2], Megumi Sageshima[1], Katsushi Tokunaga[1]

1 Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, 2 Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Japan

Abstract

The DigiTag2 assay enables analysis of a set of 96 SNPs using Kapa 2GFast HotStart DNA polymerase with a new protocol that has a total running time of about 7 hours, which is 6 hours shorter than the previous protocol. Quality parameters (conversion rate, call rate, reproducibility and concordance) were at the same levels as when genotype calls were acquired using the previous protocol. Multiplex PCR with 192 pairs of locus-specific primers was available for target preparation in the DigiTag2 assay without the optimization of reaction conditions, and quality parameters had the same levels as those acquired with 96-plex PCR. The locus-specific primers were able to achieve sufficient (concentration of target amplicon $\geq$5 nM) and specific (concentration of unexpected amplicons $<$2 nM) amplification within 2 hours, were also able to achieve detectable amplifications even when working in a 96-plex or 192-plex form. The improved DigiTag2 assay will be an efficient platform for screening an intermediate number of SNPs (tens to hundreds of sites) in the replication analysis after genome-wide association study. Moreover, highly parallel and short-acting amplification with locus-specific primers may thus facilitate widespread application to other PCR-based assays.

Introduction

Polymerase chain reaction (PCR) is a commonly used technique in molecular biology. Several previously developed methods have employed multiplexed PCR in order to analyze genomic variations such as microsatellites or short tandem repeats (STRs), single nucleotide polymorphisms (SNPs) and insertions/deletions [1–3]. Multiplexed preparation of DNA templates in a single reaction is cost-effective, saving starting materials and run-time, while requiring careful optimization of assay conditions. The optimization process is highly empirical and time consuming, and depending on the combinations of markers, may or may not lead to successful assay development. For the conventional design of multiplex PCR, optimization of reaction conditions and careful pre-selection of targets are required in order to prevent excessive off-target priming by the numerous primers in the reaction. Moreover, the risk of generating errors in multiplex PCR, such as insufficient amplification, biased amplification and considerable primer-dimer formation within primers, tends to increase roughly as the square of the number of added primer pairs [4].

There are several approaches to resolving these drawbacks, including solid-phase assay formats (glass slide arrays, microbeads), oligonucleotides containing locked nucleic acid (LNA) residues and circularized amplification. Primers immobilized on the surface of the solid phase appear to markedly increase product yield on solid supports and may avoid the need for target pre-selection with a modification to enrich the input genomic DNA via a crude solution-phase multiplex PCR [5,6]. LNA pentamers showed high priming efficiency to achieve small biased priming in multiplex PCR [7]. Circularized amplification avoids generating artifacts associated with conventional multiplex PCR where two primers are used for each target [8]. This procedure was shown to perform a 96-plex amplification of an arbitrary set of specific DNA sequences. The arrayed primer extension-based genotyping method (APEX-2) allows efficient homogeneous 640-plex DNA amplification with locus-specific primers [9]. These approaches show effective consequences for multiplex amplification, however, a small number of approaches are practically used in the field of molecular genetics, presumably due to its cost and time consuming steps in preparation.

We developed the DigiTag2 assay for multiplex SNP typing as a simple and cost effective approach by combining multiplex PCR to enrich genetic regions including the target SNPs and an oligonucleotide ligation assay to encode all of the SNP genotypes into well-designed oligonucleotides designated DNA coded numbers (DCNs) [10]. For an effective primer design for multiplex PCR, there are several important physical properties for primer sequences, including melting temperature, Gibbs energy of duplex between primer and template, and interactions between primers and PCR amplicons. The DNA polymerase enzyme used in a multiplex PCR is one of the important factors for a successful unbiased amplification.

The DigiTag2 assay is a suitable approach to analyze an intermediate number of SNPs (tens to hundreds of locus) in the replication study after genome wide association study [11–12]. However, the most time consuming step for the DigiTag2 assay in a total running time of 13 hours is multiplex PCR for target preparation (5.5 hours). Here, we report an improved protocol for the DigiTag2 assay with a short-acting multiplex PCR through the use of Kapa 2GFast HotStart DNA polymerase, which reduces total running time and increases assay throughput. In this study, we also validate the applicability of the 192-plex PCR with locus specific primers to amplify the target regions from genomic DNA, which leads to save genomic DNA samples.

Methods

DNA samples

Genomic DNA samples from 96 unrelated healthy donors were obtained from the Japan Health Science Research Resources Bank (Osaka, Japan). All donors provided written informed consent and samples were anonymized. One microgram of purified genomic DNA was dissolved in 100 µl of TE buffer (pH 8.0) (Wako, Osaka, Japan), followed by storage at $-20°C$ until use.

Primer design

A total of 192 pairs of primer were designed using the Visual OMP software version 7.1.0.0 (DNA software, Ann Arbor, MI, USA) with relatively long length (35–45-mer; average, 39.5-mer) to give amplicon sizes between 312 bp and 995 bp (average, 589 bp), each of which had an SNP site (Table S1). Prediction of DNA melting temperature was calculated using nearest-neighbor thermodynamic models. To avoid spurious amplification products, we employed a two-step protocol (denature and extension steps) using specifically designed primer pairs with an extension temperature at $68°C$. The specificity of primer sequences was verified by Blat search in order to predict its location(s) on the human genome (GRCh37), and to confirm no unexpected SNP(s) within the primer sequence. The specificity of primer pairs was verified using MFE primer software, which can predict potential amplicon(s) generated from the human genome (GRCh37, up to 5 kb in amplicon size) [13]. All oligonucleotides (de-salted, 100 pmol/µl in TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA)) were purchased from Life Technologies (Carlsbad, CA, USA), and were stored at $-20°C$.

Multiplex PCR with Kapa 2GFast HotStart DNA polymerase

Multiplex PCR mix had a final volume of 10 µl, including 10 ng of genomic DNA, 25 nM each primer, $1.5×$ KAPA2G Buffer (including 2.25 mM Mg^{2+}), an additional 2.25 mM Mg^{2+} (final concentration of Mg^{2+}: 4.5 mM), 0.2 mM dNTPs and 0.4 U of Kapa 2GFast HotStart DNA polymerase (Kapa Biosystems, Woburn, MA, USA). PCR amplification was conducted using a TGradient (Biometra, Göttingen, Germany) or PTC-225 (MJ Research, Waltham, MA, USA) as follows: $95°C$ for 3 min, followed by 40 cycles of $95°C$ for 15 s and $68°C$ for 2 min. When necessary, the fragment length of PCR products was confirmed by capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent, Santa Clara, CA, USA) in order to evaluate PCR efficiency. The total running times for multiplex PCR with Kapa 2GFast HotStart DNA polymerase using TGradient and PTC-225 were 1 h 48 min 55 s and 2 h 6 min 59 s, respectively.

Multiplex PCR with QIAGEN Multiplex PCR Kit

Multiplex PCR mix had a final volume of 10 µl, including 10 ng of genomic DNA, 25 nM each primer, $1×$ Multiplex PCR Buffer (including 3.0 mM Mg^{2+}), 0.2 mM dNTPs and HotStar-Taq DNA polymerase (QIAGEN Multiplex PCR Kit; QIAGEN, Valencia, CA, USA). PCR amplification was conducted using a TGradient or PTC-225 as follows: $95°C$ for 15 min, followed by 40 cycles of $95°C$ for 30 s and $68°C$ for 6 min. The total running times for multiplex PCR with QIAGEN Multiplex PCR Kit using TGradient and PTC-225 were 5 h 27 min 53 s and 5 h 46 min 39 s, respectively.

96-plex genotyping by the DigiTag2 assay

The DigiTag2 assay performs multiplex SNP typing by encoding all of the SNP genotypes into well-designed oligonucleotides, designated DNA coded numbers (Figure 1, DCNs: D1_i, ED-1 and ED-2) [10]. The DCNs are assigned to the target SNPs in an unconstrained manner; therefore, the DNA chips prepared to read out the types of DCNs are universally available for any type of SNP without optimization of assay conditions. The DigiTag2 assay proceeds in four steps; target preparation, encoding, labeling and detection.
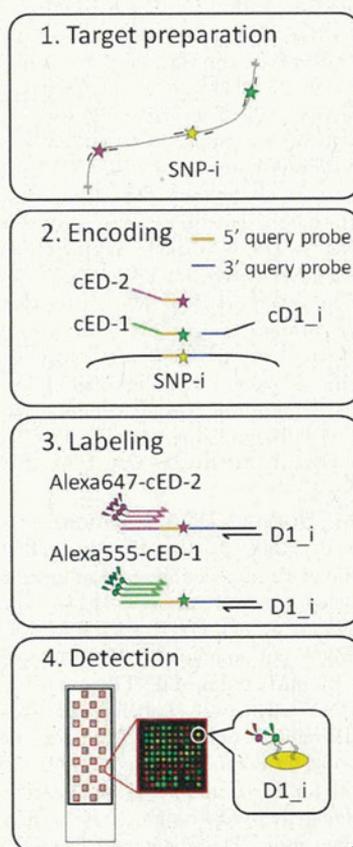
Figure 1. Schematic representation of the DigiTag2 assay. The assay has four steps: target preparation, encoding, labeling and detection. SNP genotypes are encoded into well-designed oligonucleotides, designated DNA coded numbers (DCNs: D1_i, ED-1 and ED-2). D1_i is a variable sequence assigned to each SNP. Reverse complement sequences are written by attaching the character 'c' before the sequence name.
doi:10.1371/journal.pone.0029967.g001

The encoding reactions had a final volume of 15 µl, including 0.5 µl of multiplex PCR products, 20 mM Tris-HCl, pH 7.6, 25 mM potassium acetate, 10 mM magnesium acetate, 10 mM DTT, 1 mM NAD, 0.1% Triton X-100 (1× Taq DNA ligase buffer) with 0.33 nM of each probe and 5 U Taq DNA ligase (New England BioLabs, Ipswich, MA, USA). Encoding reactions were conducted using a TGradient or PTC-225 under the following conditions: 95°C for 5 min, followed by 58°C for 15 min. The reaction was stopped by holding the temperature at 10°C.

The labeling reactions had a final volume of 12 µl, including 6 µl of ligation products, 0.5 µM each labeled primer (Alexa555-cED-1 and Alexa647-cED-2), 2.5 nM each D1 primer (D1_i), 50 mM KCl, 2 mM Mg^{2+}, 0.1 mM DTT, 0.2 mM each dNTP (N = A, G, C), 0.1 mM [^{3}H]-dTTP, 0.25 mg/ml activated salmon sperm DNA (1× Ex Taq Buffer) and 0.05 U of Ex Taq^{TM} polymerase (TaKaRa, Shiga, Japan). Labeling reactions were conducted using a TGradient or PTC-225 under the following conditions: first held at 95°C for 1 min, followed by 30 cycles of 95°C for 30 s, 55°C for 6 min and 72°C for 30 s. The reaction was stopped by holding the temperature at 10°C. Total running times for labeling using TGradient and PTC-225 were 3 h 49 min 48 s and 4 h 8 min 48 s, respectively.

In the detection step, a hybridization mixture was prepared by mixing 6.25 µl of labeling products with 8.75 µl of hybridization buffer containing 0.5× SSC, 0.1% SDS, 15% formamide, 1 mM EDTA and 3.125 fmol of hybridization control (Alexa555-labeled D1_100 and Alexa647-labeled D1_100). The hybridization control was prepared for ensuring the hybridization step. Ten microliters of hybridization mixture was applied to each block on the universal DNA chip. Hybridization was carried out for 30 min at 37°C in a hybridization oven (ThermoStat plus; Eppendorf, Ham, Germany). After hybridization, glass slides were washed in washing buffer (0.1× SSC, 0.1% SDS) by shaking at 60 rpm for 3 min. Glass slides were consecutively washed in distilled water by shaking at 60 rpm for 1 min and then dried up by centrifugation at 500× g for 1 min. Hybridization images were scanned at photomultiplier voltages of 400 V for Alexa555 and 480 V for Alexa647 using a commercially available DNA chip scanner and fluorescence image analysis was performed using commercially available software (GenePix 4000B unit and GenePix Pro 4.1 software package; Molecular Devices, Sunnyvale, CA, USA).

Labeling with Kapa 2GFast HotStart DNA polymerase

The labeling reactions with Kapa 2GFast HotStart DNA polymerase had a final volume of 12 µl, including 6 µl of ligation products, 0.5 µM each labeled primer (Alexa555-cED-1 and Alexa647-cED-2), 2.5 nM each D1 primer (D1_i), 1.5× KAPA2G Buffer (including 2.25 mM Mg^{2+}), an additional 2.25 mM Mg^{2+} (final concentration of Mg^{2+}: 4.5 mM), 0.2 mM dNTPs and 0.4 U of Kapa 2GFast HotStart DNA polymerase. Labeling reactions were conducted using a TGradient or PTC-225 under the following conditions: first held at 95°C for 1 min, followed by 30 cycles of 95°C for 15 s, 55°C for 120 s and 72°C for 5 s. The reaction was stopped by holding the temperature at 10°C. The total running times for labeling using TGradient and PTC-225 were 1 h 29 min 48 s and 1 h 48 min 34 s, respectively.

Results

Singleplex PCR using 192 pairs of locus-specific primers

Singleplex PCR was conducted under the same reaction condition with multiplex PCR using 25 ng of genomic DNA to ensure target amplicon detection and to confirm the emergence of extra bands (unexpected amplicons). Singleplex PCR with 192 pairs of locus-specific primers revealed that most of the primer pairs are able to achieve sensitive detection (concentration of target amplicon ≥5 nM) and specific amplification without extra bands (concentration of unexpected amplicons <2 nM) except for 14 pairs of primers; low sensitivity (<5 nM) for 5 pairs of primers (61, 99, 102, 189 and 191) and low specificity with extra bands (≥2 nM) for 9 pairs of primers (40, 56, 62, 70, 91, 106, 149, 173 and 174) (Figure 2 and Table S2). Five pairs among the 9 low-specific primer pairs with extra bands (62, 70, 149, 173 and 174) resulted from heteroduplex formation of target amplicons during polyacrylamide gel electrophoresis. Despite the presence of extra bands, the remaining 4 pairs of low-specific primers had a target amplicon with a detectable concentration ≥5 nM.

Validation of efficacy of 192-plex PCR by 96-plex genotyping with the DigiTag2 assay

The DigiTag2 assay enables the simultaneous analysis of 96 target SNPs in: (1) multiplex PCR with locus-specific primers to amplify target genomic regions including target SNPs; (2) multiple oligonucleotide ligation assay with locus-specific probes to determine the genotype of each SNP; and (3) hybridization to the universal DNA chip tethered with probe sequences identical to D1_i (23-mer) (Figure 1) [10]. The validity of 192-plex PCR was assessed with 96 individual DNAs (population control samples) by comparing two sets of 96-plex genotype calls acquired from 96-plex PCR with those from 192-plex PCR (Table 1).

Conversion rate shows the proportion of successfully genotyped SNPs with fewer than 3 undetected samples after excluding low-quality genotyping data, which had more than 5 undetected SNPs in a total of 96 SNPs. However, the composition of failed SNPs in genotyping was not identical, and the conversion rate showed no differences between 192-plex PCR and 96-plex PCR. For the 1st set of 96 SNPs, 7 SNPs among 10 failed SNPs were matched between 192-plex PCR and 96-plex PCR, and for the 2nd set, 8 SNPs among the 9 failed SNPs were matched. The average call rate for successfully genotyped SNPs was over 99.79% for both sets of 96-plex genotyping, even if 192-plex PCR products were adopted for target preparation. Reproducibility was determined by independent genotyping with 96 individuals twice. As a consequence, four discordant genotype calls were observed in the duplicated genotyping data. Concordance of genotype calls between 192-plex PCR and 96-plex PCR was determined using 6,290 genotype calls for the 1st set and 7,884 genotype calls for the 2nd set. Consequently, 14,171 out of 14,174 genotype calls were matched by comparison with 83 SNPs for the 1st set and 86 SNPs for the 2nd set. In total, 3 discordant genotype calls were observed (Figure 3).

Short-acting multiplex PCR by use of Kapa 2GFast HotStart DNA polymerase

Kapa 2GFast HotStart DNA polymerase was employed to perform multiplex PCR with the locus-specific primers for target preparation in genotyping with the DigiTag2 assay. To optimize reaction conditions with Kapa 2GFast HotStart DNA polymerase, singleplex PCR was conducted using 25 ng of genomic DNA with three randomly chosen pairs of locus-specific primers. The designed amplicon sizes for the three pairs of primers were 501 bp, 671 bp and 492 bp. We performed singleplex PCR using a two-step protocol (denature and extension steps) with varied extension periods (15 s, 30 s, 60 s and 120 s) and with varied Mg^{2+} concentrations (3.0 mM and 4.5 mM) (Figure 4). The most sensitive detection and highest levels of amplification for the three
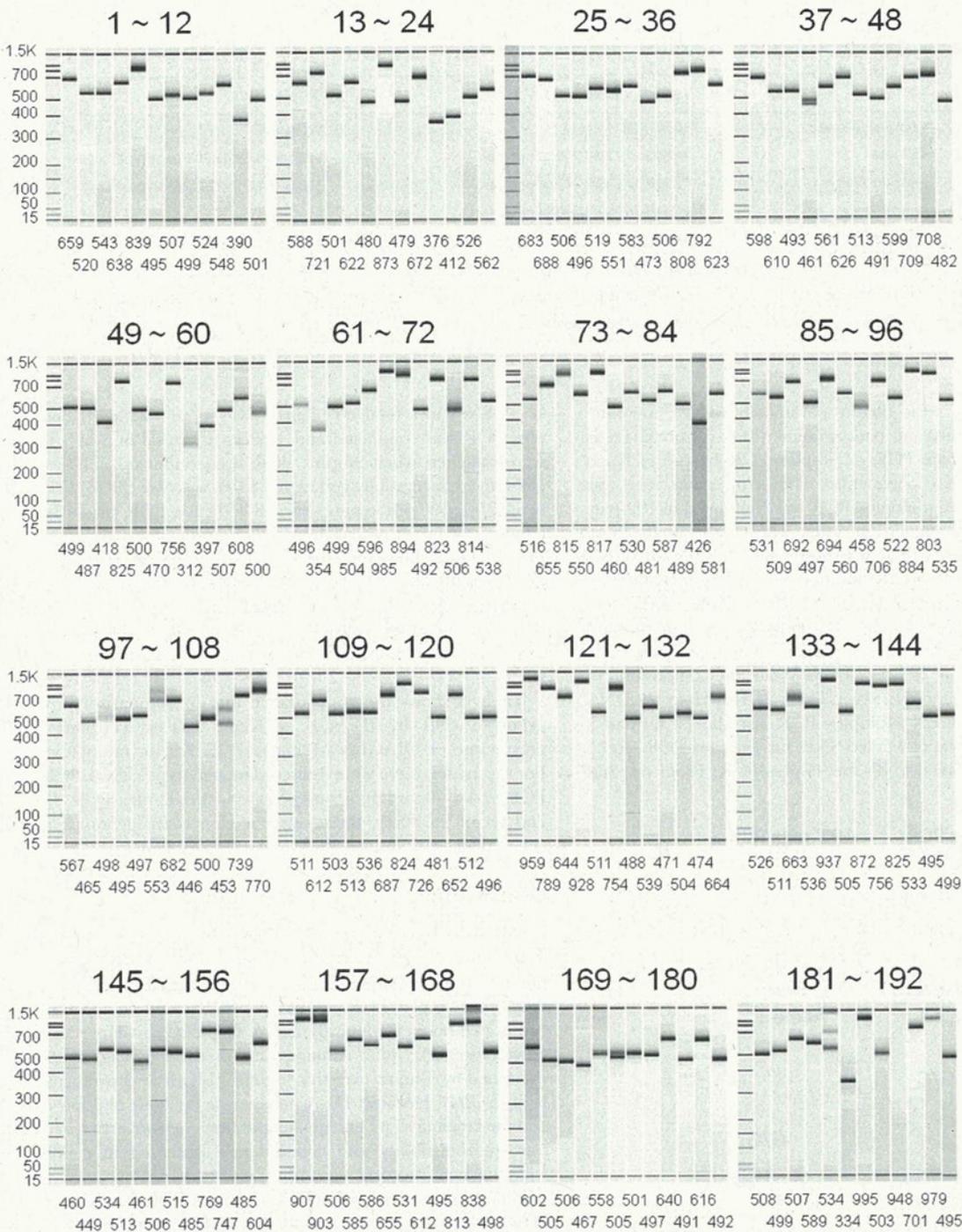
Figure 2. Electropherogram of singleplex PCR products with 192 pairs of locus-specific primers. The designed amplicon size is depicted below each lane.
doi:10.1371/journal.pone.0029967.g002

pairs of primers were observed with 120 s for the extension period and 4.5 mM for the Mg^{2+} concentration. The total running time for multiplex PCR with locus-specific primers was less than 2 hours, which is about 3 h 30 min shorter than the previous protocol (see MATERIALS AND METHODS).

The total running time of the DigiTag2 assay was markedly reduced when the labeling step was also conducted using Kapa 2GFast HotStart DNA polymerase instead of *Ex Taq* polymerase. When the DigiTag2 assay was conducted with Kapa 2GFast HotStart DNA polymerase for multiplex PCR and labeling step, the total running time of the assay was about 7 hours, which is about 6 hours shorter than the previously used protocol in combination with QIAGEN Multiplex PCR Kit for multiplex PCR and *Ex Taq* polymerase for the labeling step.

Table 1. Validation of efficacy of 192-plex PCR by 96-plex genotyping.

		192-plex PCR	96-plex PCR
1st set	Conversion rate	86/96 SNP	86/96 SNP
	Call rate	99.84% (7,728/7,740 genotype)	99.81% (6,695/6,708 genotype)
	reproducibility	99.99% (7,288/7,289 genotype)	100% (6,121/6,121 genotype)
	concordance	99.98% (6,289/6,290 genotype)	
2nd set	Conversion rate	87/96 SNP	87/96 SNP
	Call rate	99.79% (8,074/8,091 genotype)	99.79% (8,161/8,178 genotype)
	reproducibility	99.97% (7,792/7,794 genotype)	99.99% (7,712/7,713 genotype)
	concordance	99.97% (7,882/7,884 genotype)	

doi:10.1371/journal.pone.0029967.t001

Table 2 summarizes the quality parameters (conversion rate, call rate, reproducibility and concordance) when genotyping was conducted with 192-plex PCR or 96-plex PCR by use of Kapa 2GFast HotStart DNA polymerase. The conversion rate was slightly decreased when multiplex PCR was conducted in 192-plex form. However, the conversion rates were better than those observed when multiplex PCR was conducted with the QIAGEN Multiplex PCR Kit. The composition of failed SNPs in genotyping was not consistent for the 1st set of 96 SNPs, in which 4 SNPs were matched between 192-plex PCR and 96-plex PCR. For the 2nd set, a total of 8 failed SNPs in the 96-plex PCR were completely matched to those in the 192-plex PCR. When the composition of failed SNPs were compared between Kapa 2GFast HotStart DNA polymerase and QIAGEN Multiplex PCR Kit, the 1st set had 5 matched SNPs in a total of 8 failed SNPs for 192-plex PCR, and 4 matched SNPs in 5 failed SNPs for 96-plex PCR. From the 2nd set, 5 SNPs in a total of 9 failed SNPs were matched when 192-plex PCR was conducted and 4 SNPs in a total of 8 failed SNPs were matched when 96-plex PCR was conducted. The average call rate for successfully genotyped SNPs was over 99.76% for both sets of 96-plex genotyping, even if 192-plex PCR products were adopted for target preparation. The reproducibility was 100% for the 2nd set; however, three discordant genotype calls were observed for the 1st set. With regard to the concordance of genotype calls between 96-plex PCR and 192-plex PCR, only one discordant genotype call was observed in the comparison for the 1st set, and no discordant genotype calls were observed in the 2nd set.

Table 3 shows the concordance rate in comparison with the genotype calls by the use of Kapa 2GFast HotStart DNA polymerase or QIAGEN Multiplex PCR Kit for multiplex PCR. For the 1st set, there were 4 discordant genotype calls with 96-plex PCR and 8 discordant genotype calls with 192-plex PCR. For the 2nd set of 96 SNPs, there was one discordant genotype call in genotyping with 96-plex PCR and 192-plex PCR.
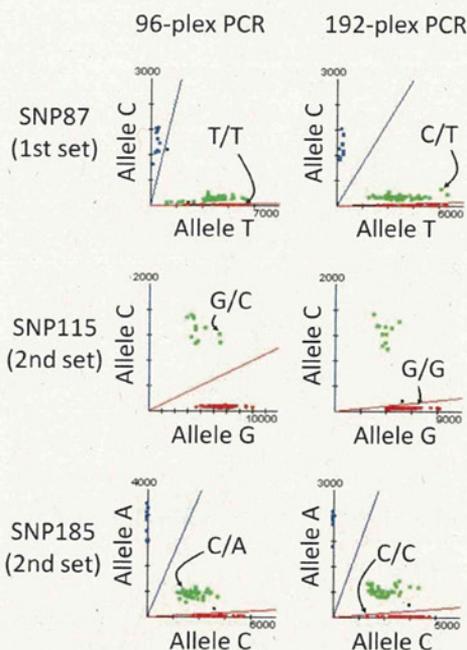
Discussion

The locus specific primers sufficiently worked in a multiplex form under the same reaction conditions without any optimization processes, either 96-plex PCR or 192-plex PCR. We also found that either 96-plex PCR or 192-plex PCR could be accomplished within two hours through the use of Kapa 2GFast HotStart DNA polymerase. The total running time of the DigiTag2 assay was shortened by 6 hours over the original 13-hour long protocol using Kapa 2GFast HotStart DNA polymerase for both multiplex PCR and the labeling step. The quality parameters (conversion rate, call rate, reproducibility and concordance) observed in genotyping with the new protocol were the same as those observed in the original protocol using QIAGEN Multiplex PCR Kit for multiplex PCR and *Ex Taq* polymerase for the labeling step. The DigiTag2 assay worked with a conversion rate of over 93.2% (179 / 192 SNPs), average call rate of over 99.80% (16,789/16,823 genotypes) and reproducibility of over 99.99% (16,135/16,136 genotypes) using 96-plex PCR under the new protocol. The composition of successfully genotyped SNPs was different when the genotype calls were acquired using the different polymerases (Kapa 2GFast HotStart DNA polymerase and QIAGEN Multiplex PCR Kit), which would result from a varying amplification bias in multiplex PCR. We also found that 192-plex PCR with locus-specific primers worked in 96-plex genotyping with the DigiTag2 assay, giving the same quality parameter data as those observed in genotyping with 96-plex PCR. However, the

96-plex PCR   192-plex PCR

SNP87 (1st set)

SNP115 (2nd set)

SNP185 (2nd set)

Figure 3. Scatter plots for three SNPs with 3 discordant genotypes. Scatter plots in genotyping with 192-plex PCR and 96-plex PCR are depicted side-by-side. The genotypes of discordant samples are indicated in the scatter plots by arrows.
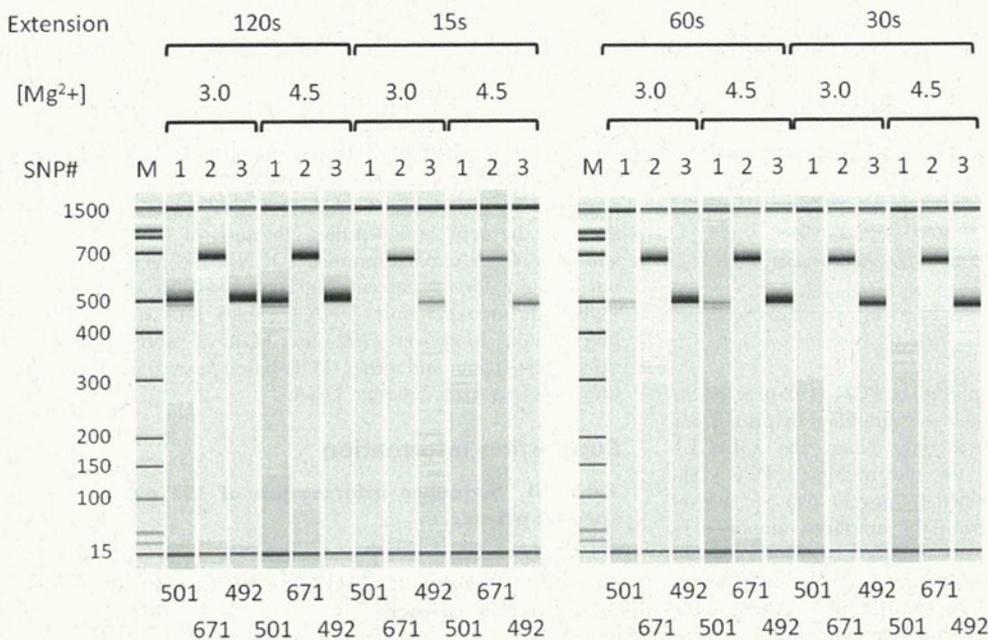doi:10.1371/journal.pone.0029967.g003

Figure 4. Electropherogram of singleplex PCR products using Kapa 2GFast HotStart DNA polymerase. Singleplex PCR was performed with varied extension periods (15 s, 30 s, 60 s and 120 s) and with varied Mg^{2+} concentrations (3.0 mM and 4.5 mM) using three pairs of locus-specific primers. The designed amplicon size is depicted below each lane.
doi:10.1371/journal.pone.0029967.g004

composition of successfully genotyped SNPs was not consistent between 192-plex PCR and 96-plex PCR, which may be explained by changing the interactions between primer pairs in 192-plex PCR and in 96-plex PCR. The composition of successful SNPs was not consistent when using different polymerases or multiplex systems in the multiplex PCR, which casts some shadows on the reliability of the assay. Regardless of the existing shadows, indeed, 96-plex and 192-plex PCR work with a high conversion rate in genotyping over 93.2%. To clear the existing shadows, it is necessary to continuously accumulate genotyping data.

In this study, fifteen discordant genotype calls were in total observed in the comparison of genotype calls with: i) duplicated genotyping data; ii) genotyping data by use of 192-plex PCR and 96-plex PCR; and iii) genotyping data with different types of polymerases (Table S3). Table S3 shows the genotype calls acquired 8 times under different conditions. All fifteen discordant genotype calls were analyzed with direct sequencing, of which 13 genotype calls were determined. In 8 of 15 discordant genotype

calls, the genotype calls were completely different depending on the type of polymerase. The genotype calls acquired using Kapa 2GFast HotStart DNA polymerase were 100% concordant (6 of 6) with those acquired by direct sequencing. This suggests that SNP allelic bias in PCR amplification readily occurred with the QIAGEN Multiplex PCR Kit; however, the error rate in genotyping was only 0.04% (6 out of 14,886 genotypes). The remaining 7 discordant genotype calls were randomly observed in 1 out of 8 different conditions. This shows that the random error rates were almost equal in the genotype data acquired with both types of polymerases (4 out of 62,227 genotypes for QIAGEN Multiplex PCR Kit and 3 out of 66,008 genotypes for Kapa 2GFast HotStart DNA polymerase).

Among the five low-sensitivity primer pairs found on singleplex PCR (61, 99, 102, 189 and 191), no amplicons were detected by primer pair 189 and low concentrations (<5 nM) of amplicon were detected by the 4 other primer pairs (Table S2). Therefore, the SNP189 failed in genotyping, independently of the type of

Table 2. Validation of efficacy of 192-plex and 96-plex PCR with Kapa 2GFast HotStart DNA polymerase.

		192-plex PCR	96-plex PCR
1st set	Conversion rate	88/96 SNP	91/96 SNP
	Call rate	99.84% (8,259/8,272 genotype)	99.76% (8,443/8,463 genotype)
	reproducibility	99.97% (8,069/8,071 genotype)	99.99% (8,339/8,340 genotype)
	concordance	99.99% (7,982/7,983 genotype)	
2nd set	Conversion rate	87/96 SNP	88/96 SNP
	Call rate	99.91% (8,171/8,178 genotype)	99.83% (8,346/8,360 genotype)
	reproducibility	100% (7,705/7,705 genotype)	100% (7,796/7,796 genotype)
	concordance	100% (8,161/8,161 genotype)	

doi:10.1371/journal.pone.0029967.t002

Table 3. Concordance of genotype calls between Kapa 2GFast HotStart DNA polymerase and QIAGEN Multiplex PCR Kit.

		Kapa 2G	QIAGEN
1st set	96-plex PCR	99.94% (6,513/6,517 genotype)	
	192-plex PCR	99.89% (7,441/7,449 genotype)	
2nd set	96-plex PCR	99.99% (7,778/7,779 genotype)	
	192-plex PCR	99.99% (7,700/7,701 genotype)	

doi:10.1371/journal.pone.0029967.t003

polymerase and multiplicity in multiplex PCR (192-plex or 96-plex). However, the SNP191, which was amplified by primer pair 191, was successfully genotyped only when the QIAGEN Multiplex PCR Kit was used for the multiplex PCR. The concentration of amplicon amplified by primer pair 99 was the same as the 2.8 nM observed with the amplicon amplified by primer pair 191. SNP99, which was amplified by primer pair 99, was successfully genotyped independently of polymerase type and multiplicity in multiplex PCR (192-plex or 96-plex). These results suggest that the sensitivity in genotyping with Kapa 2GFast HotStart DNA polymerase was lower than the previously used protocol with QIAGEN Multiplex PCR Kit. These results would be explained by a biased amplification with the shortened protocol using Kapa 2GFast HotStart DNA polymerase, which tends to lead to a consequent biased genotyping. However, the investigated number of primer pairs would not be sufficient to decide the sensitivity in genotyping; therefore, it is necessary to continuously accumulate genotyping data. As the investigated number of primer pairs was only 192 (384 primers) in this study, melting temperature of each primer and the number of potential amplicons predicted by the MFE primer software were strongly associated with low sensitivity and low specificity in an amplification, respectively (multiple regression analysis, $P = 1.26 \times 10^{-37}$ and $P = 1.52 \times 10^{-21}$, respectively).

Through the use of Kapa 2GFast HotStart DNA polymerase, the genotype calls for 96 SNPs can be acquired in about 7 hours by the DigiTag2 assay. The genotyping platform with high conversion rate plays an important role for the replication studies to identify the disease associated genes from candidate loci found in the GWAS (genome-wide association study). The DigiTag2 assay with an improved protocol will be an efficient platform for screening an intermediate number of SNPs (tens to hundreds of sites) in the replication studies. Because of limitations in the variation of DNA coded numbers (DCNs), 192-plex genotyping is not available for the current DigiTag2 assay. However, 192-plex PCR can save genomic DNA samples and time for target preparation. Moreover, 192-plex PCR is also available for direct-sequencing and other PCR-based assays to amplify the target regions from genomic DNA.

Supporting Information

Table S1 Sequence information of 192 pairs of locus specific primer.
(XLSX)

Table S2 Results of singleplex PCR with 192 pairs of locus specific primer.
(XLSX)

Table S3 The 15 discordant genotype calls in 8 different conditions.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: NN KT. Performed the experiments: YM MS. Analyzed the data: NN YM MS. Contributed reagents/materials/analysis tools: NN YM MS. Wrote the paper: NN KT.

References

1. Deter J, Gala M, Charbonnel N, Cosson JF (2009) Characterization and PCR multiplexing of polymorphic microsatellite loci in the whipworm *Trichuris arvicolae*, parasite of arvicoline rodents and their cross-species utilization in *T. muris*, parasite of murines. Mol Biochem Parasitol 167: 144–146.
2. Hosseini-Maaf B, Hellberg Å, Chester MA, Olsson ML (2007) An extensive polymerase chain reaction–allele-specific polymorphism strategy for clinical ABO blood group genotyping that avoids potential errors caused by null, subgroup, and hybrid alleles. Transfusion 47: 2110–2125.
3. Goguet de la Salmonière YO, Kim CC, Tsolaki AG, Pym AS, Siegrist MS, et al. (2004) High-throughput method for detecting genomic-deletion polymorphisms. J Clin Microbiol 42: 2913–2918.
4. Landegren U, Nilsson M (1997) Locked on target: strategies for future gene diagnostics. Ann Med 29: 585–590.
5. Pemov A, Modi H, Chandler DP, Bavykin S (2005) DNA analysis with multiplex microarray-enhanced PCR. Nucleic Acids Res 33: e11.
6. Meuzelaar LS, Lancaster O, Pasche JP, Kopal G, Brookes AJ (2007) MegaPlex PCR: a strategy for multiplex amplification. Nat Methods 4: 835–837.
7. Sun Z, Chen Z, Hou X, Li S, Zhu H, et al. (2008) Locked nucleic acid pentamers as universal PCR primers for genomic DNA amplification. Plos One 3: e3701.
8. Dahl F, Gullberg M, Stenberg J, Landegren U, Nilsson M (2005) Multiplex amplification enabled by selective circularization of large sets of genomic DNA fragments. Nucleic Acids Res 33: e71.
9. Krjutškov K, Andreson R, Mägi R, Nikopensius T, Khrunin A, et al. (2008) Development of a single tube 640-plex genotyping method for detection of nucleic acid variations on microarrays. Nucleic Acids Res 36: e75.
10. Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 364: 78–85.
11. Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, et al. (2009) Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41: 1105–1109.
12. Miyagawa T, Kawashima M, Nishida N, Ohashi J, Kimura R, et al. (2009) Variant between CPT1B and CHKB associated with susceptibility to narcolepsy. Nat Genet 40: 1324–1328.
13. Qu W, Shen Z, Zhao D, Yang Y, Zhang C (2009) MFEprimer: multiple factor evaluation of the specificity of PCR primers. Bioinformatics 25: 276–278.

PLoS one

Genome-Wide Association Study Confirming Association of HLA-DP with Protection against Chronic Hepatitis B and Viral Clearance in Japanese and Korean

Nao Nishida[1,2]*, Hiromi Sawai[2], Kentaro Matsuura[3], Masaya Sugiyama[1], Sang Hoon Ahn[4], Jun Yong Park[4], Shuhei Hige[5], Jong-Hon Kang[6], Kazuyuki Suzuki[7], Masayuki Kurosaki[8], Yasuhiro Asahina[8], Satoshi Mochida[9], Masaaki Watanabe[10], Eiji Tanaka[11], Masao Honda[12], Shuichi Kaneko[12], Etsuro Orito[13], Yoshito Itoh[14], Eiji Mita[15], Akihiro Tamori[16], Yoshikazu Murawaki[17], Yoichi Hiasa[18], Isao Sakaida[19], Masaaki Korenaga[20], Keisuke Hino[20], Tatsuya Ide[21], Minae Kawashima[2], Yoriko Mawatari[1,2], Megumi Sageshima[2], Yuko Ogasawara[2], Asako Koike[22], Namiki Izumi[8], Kwang-Hyub Han[4], Yasuhito Tanaka[3], Katsushi Tokunaga[2], Masashi Mizokami[1]

1 Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Chiba, Japan, 2 Department of Human Genetics, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, 3 Department of Virology and Liver Unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Aichi, Japan, 4 Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea, 5 Department of Internal Medicine, Hokkaido University Graduate School of Medicine, Sapporo, Japan, 6 Department of Internal Medicine, Teine Keijinkai Hospital, Sapporo, Japan, 7 Department of Gastroenterology and Hepatology, Iwate Medical University, Morioka, Japan, 8 Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan, 9 Division of Gastroenterology and Hepatology, Saitama Medical University, Saitama, Japan, 10 Department of Gastroenterology, Kitasato University School of Medicine, Sagamihara, Kanagawa, Japan, 11 Department of Medicine, Shinshu University School of Medicine, Matsumoto, Japan, 12 Department of Gastroenterology, Kanazawa University Graduate School of Medicine, Kanazawa, Japan, 13 Department of Gastroenterology, Nagoya Daini Red Cross Hospital, Nagoya, Japan, 14 Molecular Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, Kyoto, Japan, 15 Department of Gastroenterology and Hepatology, National Hospital Organization Osaka National Hospital, Osaka, Japan, 16 Department of Hepatology, Osaka City University Graduate School of Medicine, Osaka, Japan, 17 Second Department of Internal Medicine, Faculty of Medicine, Tottori University, Yonago, Japan, 18 Department of Gastroenterology and Metabology, Ehime University Graduate School of Medicine, Ehime, Japan, 19 Gastroenterology and Hepatology, Yamaguchi University Graduate School of Medicine, Yamaguchi, Japan, 20 Division of Hepatology and Pancreatology, Kawasaki Medical College, Kurashiki, Japan, 21 Division of Gastroenterology, Department of Medicine, Kurume University School of Medicine, Fukuoka, Japan, 22 Central Research Laboratory, Hitachi Ltd., Kokubunji, Tokyo, Japan

Abstract

Hepatitis B virus (HBV) infection can lead to serious liver diseases, including liver cirrhosis (LC) and hepatocellular carcinoma (HCC); however, about 85–90% of infected individuals become inactive carriers with sustained biochemical remission and very low risk of LC or HCC. To identify host genetic factors contributing to HBV clearance, we conducted genome-wide association studies (GWAS) and replication analysis using samples from HBV carriers and spontaneously HBV-resolved Japanese and Korean individuals. Association analysis in the Japanese and Korean data identified the *HLA-DPA1* and *HLA-DPB1* genes with $P_{meta} = 1.89 \times 10^{-12}$ for rs3077 and $P_{meta} = 9.69 \times 10^{-10}$ for rs9277542. We also found that the *HLA-DPA1* and *HLA-DPB1* genes were significantly associated with protective effects against chronic hepatitis B (CHB) in Japanese, Korean and other Asian populations, including Chinese and Thai individuals ($P_{meta} = 4.40 \times 10^{-19}$ for rs3077 and $P_{meta} = 1.28 \times 10^{-15}$ for rs9277542). These results suggest that the associations between the *HLA-DP* locus and the protective effects against persistent HBV infection and with clearance of HBV were replicated widely in East Asian populations; however, there are no reports of GWAS in Caucasian or African populations. Based on the GWAS in this study, there were no significant SNPs associated with HCC development. To clarify the pathogenesis of CHB and the mechanisms of HBV clearance, further studies are necessary, including functional analyses of the HLA-DP molecule.

Introduction

Overall, one-third of the world's population (2.2 billion) is infected with hepatitis B virus (HBV), and about 15% of these are chronic carriers. About 75% of the chronic carriers live in the east-south Asia and east pacific area, and there are 1.3–1.5 million chronic carriers living in Japan [1]. Of chronic carriers, 10–15% develop liver cirrhosis (LC), liver failure and hepatocellular carcinoma (HCC), and the remaining individuals eventually achieve a state of nonreplicative infection, resulting in hepatitis B surface antigen (HBsAg) negative and hepatitis B core antibody (anti-HBc) positive, i.e. HBV-resolved individuals [2–3]. In Japan, although the major route of HBV transmission was perinatal transmission and horizontal transmission in early childhood, infant HBV carriers have successfully been reduced since 1986 through a selective vaccination policy by the Japanese government [4–7]. However, the prevalence of HBV genotype A in acute HBV (AHB) infection has increased markedly since 2000, reaching approximately 52% in 2008 due to the lack of a universal HB vaccination, and around 10% of AHB cases could be persistent infection [8–9]. Viral factors, as well as host factors, are thought to be associated with persistent HB infection.

In 2009, significant associations between chronic hepatitis B (CHB) and a region including HLA-DPA1 and HLA-DPB1 were identified using 786 Japanese individuals having CHB and 2,201 control individuals through a two-stage genome-wide association study (GWAS) [10]. The same group was also subjected to a second GWAS using a total of 2,667 Japanese persistent HBV infection cases and 6,496 controls, which confirmed significant associations between the HLA-DP locus and CHB, in addition to associations with another two SNPs located in the genetic region including the HLA-DQ gene [11]. The associations between HLA-DP variants with HBV infection were replicated in other Asian populations, including Thai and Han Chinese individuals [10,12–13]. With regard to HBV clearance, the association between the human leukocyte antigen (HLA) class II allele and clearance of HBV was confirmed by the candidate gene approach in African, Caucasian and Asian populations [14–18]. However, in a previous GWAS using samples of Japanese CHB and control individuals, the clinical data on HBV exposure in the control individuals were unknown, and this may have led to bias. Moreover, there have been no reports of GWAS using samples from HBV carriers and HBV-resolved individuals to identify host genetic factors associated with HBV clearance other than HLA class II molecules.

Here, we performed a GWAS using samples from Japanese HBV carriers, healthy controls and spontaneously HBV-resolved individuals in order to confirm or identify the host genetic factors related to CHB and viral clearance. In the subsequent replication analysis, we validated the associated SNPs in the GWAS using two independent sets of Japanese and Korean individuals. In our study, healthy controls were randomly selected with clinically no evidence of HBV exposure, therefore, HBV-resolved individuals were prepared to clearly identify the host genetic factors related with CHB or HBV clearance.

Results

Protective Effects Against Chronic Hepatitis B in Japanese and Korean Individuals

In this study, we conducted a GWAS using samples from 181 Japanese HBV carriers (including asymptomatic carriers (ASC), CHB cases, LC cases and HCC cases, based on the criteria described in Materials and Methods) and 184 healthy controls in order to identify the host genetic factors related to progression of CHB. All samples were genotyped using a genome-wide SNP typing array (Affymetrix Genome-Wide Human SNP Array 6.0 for 900 K SNPs). Figure 1a shows a genome-wide view of the single point association data based on allele frequencies using the SNPs that met the following filtering criteria: (i) SNP call rate $\geq$95%; (ii) minor allele frequency (MAF) $\geq$1% for HBV carriers and healthy controls; and (iii) no deviation from Hardy-Weinberg equilibrium (HWE) $P\geq0.001$ in healthy controls. We identified significant associations of protective effects against CHB with two SNPs (rs3077 and rs9277542) using the allele frequency model, both of which are located in the 3' UTR of HLA-DPA1 and in the sixth exon of HLA-DPB1, respectively (rs3077, $P = 1.14\times10^{-7}$, and rs9277542, $P = 5.32\times10^{-8}$, respectively). The association for rs9277542 reached a genome-wide level of significance in the GWAS panel (Bonferroni criterion $P<8.36\times10^{-8}$ (0.05/597,789)).

In order to validate the results of GWAS, a total of 32 SNPs, including the associated two SNPs (rs3077 and rs9277542), were selected for replication in two independent sets of HBV carriers and healthy controls (replication-1:256 Japanese HBV carriers and 236 Japanese healthy controls; and replication-2:344 Korean HBV carriers and 151 Korean healthy controls; Table 1). The associations for the original significant SNP (rs9277542) and marginal SNP (rs3077) on GWAS were replicated in both replication sets [replication-1 (Japanese); rs3077, $P = 2.70\times10^{-8}$, OR = 0.48 and rs9277542, $P = 3.33\times10^{-6}$, OR = 0.54; replication-2 (Korean); rs3077, $P = 2.08\times10^{-6}$, OR = 0.47 and rs9277542, $P = 8.29\times10^{-5}$, OR = 0.54, Table 2]. We conducted meta-analysis to combine these studies using the DerSimonian Laird method (random effects model) to incorporate variation among studies. As shown in Table 2, the odds ratios were quite similar across the three studies (GWAS and two replication studies) and no heterogeneity was observed ($P_{het} = 0.80$ for rs3077 and 0.40 for rs9277542). P_{meta} values were 4.40×10^{-19} for rs3077 (OR = 0.46, 95% confidence interval (CI) = 0.39–0.54), and 1.28×10^{-15} for rs9277542 (OR = 0.50, 95% CI = 0.43–0.60). Among the remaining 30 SNPs in the replication study, 27 SNPs were successfully genotyped by the DigiTag2 assay with SNP call rate $\geq$ 95% and HWE p-value $\geq$ 0.01. Two SNPs (rs9276431 and rs7768538), located in the genetic region including the HLA-DQ gene, were marginally replicated in the two sets of HBV carriers and healthy controls with Mantel-Haenszel P values of 2.80×10^{-7} (OR = 0.56, 95% CI = 0.45–0.70) and 1.09×10^{-7} (OR = 0.53, 95% CI = 0.42–0.67), respectively, when using additive, two-tailed Cochran Mantel-Haenszel (CMH) fixed-effects model with no evidence of heterogeneity ($P_{het} = 0.67$ for rs9276431 and 0.70 for rs7768538) (Table S1).

Meta-analysis using the random effects model across 6 independent studies, including 5 additional published data, showed $P_{meta} = 3.94\times10^{-45}$, OR = 0.55 for rs3077, $P_{meta} = 1.74\times10^{-21}$, OR = 0.61 for rs9277535 and $P_{meta} = 1.69\times10^{-15}$, OR = 0.51 for rs9277542, with the SNP rs9277535 being located about 4-kb upstream from rs9277542 and showing strong linkage disequilibrium of $r^2 = 0.955$ on the HapMap JPT (Table S2). As shown in Table S2, the odds ratio was very similar among the 6 studies, and heterogeneity was negligible with $P_{het} > 0.01$.

Moreover, based on GWAS using samples from 94 chronic HBV carriers with LC or HCC and 87 chronic HBV carriers without LC and HCC, we found no significant SNPs associated with CHB progression (Figure S1).
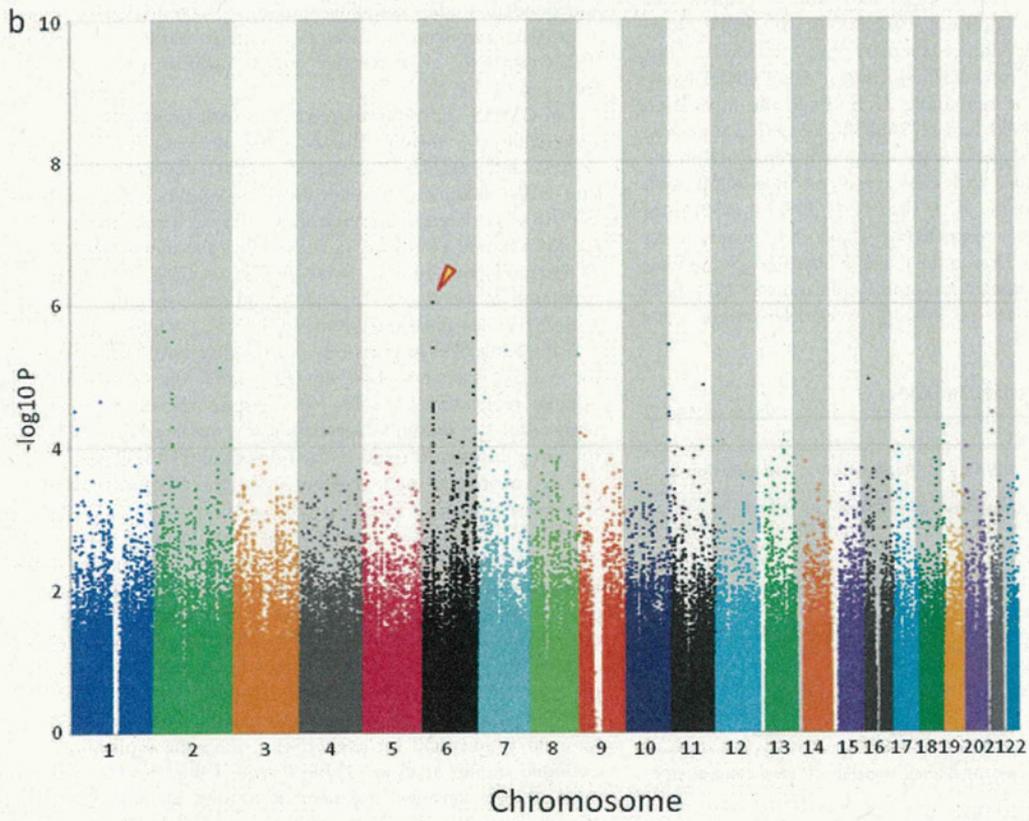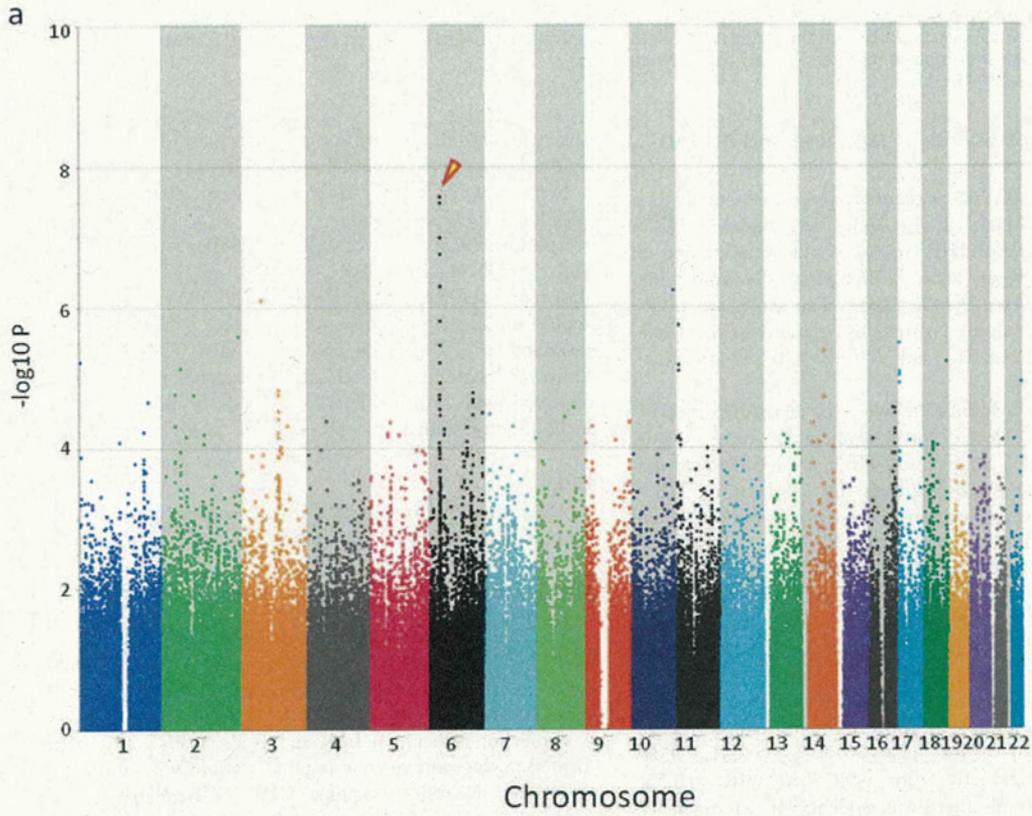
Figure 1. Results of genome-wide association studies. a) HBV carriers and healthy controls, and b) HBV carriers and HBV-resolved individuals were compared. *P* values were calculated by chi-squared test for allele frequencies. Dots with arrows on chromosome 6 show strong associations with protective effects against persistent HB infection and with HBV clearance. doi:10.1371/journal.pone.0039175.g001

Clearance of Hepatitis B virus in Japanese and Korean Individuals

We also conducted a GWAS to identify the host genetic factors related to clearance of HBV in the above 181 Japanese HBV carriers and 185 Japanese HBV-resolved individuals using a genome-wide SNP typing array (Affymetrix Genome-Wide Human SNP Array 6.0 for 900 K SNPs). The same two SNPs (rs3077 and rs9277542) showed strong associations in the allele frequency model ($P = 9.24 \times 10^{-7}$ and $P = 3.15 \times 10^{-5}$) with clearance of HBV (Figure 1b).

The above 32 SNPs, including the two associated SNPs (rs3077 and rs9277542), were selected for a replication study in two independent sets of HBV carriers and HBV resolved individuals (replication-1:256 Japanese HBV carriers and 150 Japanese HBV resolved individuals; and replication-2:344 Korean HBV carriers and 106 Korean HBV resolved individuals; Table 1). All 32 SNPs were genotyped using the DigiTag2 assay and 29 of 32 SNPs were successfully genotyped (Table S3). The associations of the original SNPs were replicated in both replication sets [replication-1 (Japanese): rs3077, $P = 3.32 \times 10^{-2}$, OR $= 0.72$ and rs9277542, $P = 1.25 \times 10^{-2}$, OR $= 0.68$; replication-2 (Korean): rs3077, $P = 2.35 \times 10^{-7}$, OR $= 0.41$ and rs9277542, $P = 4.97 \times 10^{-6}$, OR $= 0.46$; Table 3]. Meta-analysis using random effects model showed $P_{meta} = 1.56 \times 10^{-4}$ for rs3077 (OR $= 0.51$, 95% CI $= 0.36$–0.72), and 5.91×10^{-7} for rs9277542 (OR $= 0.55$, 95% CI $= 0.43$–0.69). While there was evidence of heterogeneity between these studies for rs3077 ($P_{het} = 0.03$) and no evidence for rs9277542 ($P_{het} = 0.19$), significant associations with HBV clearance were observed with Mantel-Haenszel $P_{meta} = 3.28 \times 10^{-12}$ for rs3077 and 1.42×10^{-10} for rs9277542, when using CMH fixed-effects model. Among the remaining 27 SNPs in the replication study, two SNPs (rs9276431 and rs7768538), located in a genetic region including *HLA-DQ* gene, were marginally replicated in the two sets of HBV carriers and HBV resolved individuals with Mantel-Haenszel *P* values of 2.10×10^{-5} (OR $= 0.59$) and 1.10×10^{-5} (OR $= 0.56$), respectively (Table S3), when using CMH fixed-effect model. Due to the existing heterogeneity among three groups (GWAS, Replication-1 and Replication-2) ($P_{het} = 0.03$ for rs9276431 and 0.04 for rs7768538), weak associations were

observed with $P_{meta} = 0.03$ for rs9276431 and 0.02 for rs7768538 by the random effects model meta-analysis.

Meta-analysis across 6 independent studies, including 5 additional published data, showed $P_{meta} = 1.48 \times 10^{-9}$, OR $= 0.60$ for rs3077, $P_{meta} = 1.08 \times 10^{-17}$, OR $= 0.66$ for rs9277535 and $P_{meta} = 5.14 \times 10^{-5}$, OR $= 0.55$ for rs9277542 (Table S4). As shown in Table S4, the OR for the rs9277535 and rs9277542 were similar among the 6 independent studies, and heterogeneity was negligible ($P_{het} = 0.03$ for rs9277535 and 0.14 for rs9277542). However, significant level of heterogeneity for rs3077 was observed with $P_{het} = 9.57 \times 10^{-6}$ across 5 independent studies, including our study.

URLs

The results of the present GWAS are registered at a public database: https://gwas.lifesciencedb.jp/cgi-bin/gwasdb/gwas_top.cgi.

Discussion

The recent genome-wide association study showed that the SNPs located in a genetic region including *HLA-DPA1* and *HLA-DPB1* genes were associated with chronic HBV infection in the Japanese and Thai population [10,11]. In this study, we confirmed a significant association between SNPs (rs3077 and rs9277542) located in the same genetic region as *HLA-DPA1* and *HLA-DPB1* and protective effects against CHB in Korean and Japanese individuals. Mata-analysis using the random effects model across 6 independent studies including our study suggested that, widely in East Asian populations, variants in antigen binding sites of *HLA-DP* contribute to protective effects against persistent HBV infection (Table S2).

On GWAS and replication analysis with Japanese and Korean individuals, we identified associations between the same SNPs (rs3077 and rs9277542) in the *HLA-DPA1* and *HLA-DPB1* genes and HBV clearance; however, no new candidate SNPs from the GWAS were detected on replication analysis (Table S3). When the data of reference#18 was excluded from the meta-analysis across 6 independent studies, heterogeneity among 4 studies was estimated to be $P_{het} = 0.15$ and significant association of rs3077 with HBV clearance was observed with $P_{meta} = 5.88 \times 10^{-24}$, OR $= 0.56$ (Table S4). In our study, a negligible level of heterogeneity for rs3077 was also observed ($P_{het} = 0.03$) on meta-analysis by adding replication-1 (Table 3). Despite the heterogeneity in replication-1, a marginal association was observed for rs3077 with the same downward trend in the odds ratio ($P = 3.32 \times 10^{-2}$, OR $= 0.72$). Moreover, meta-analysis using GWAS and replication-2 showed significant association of $P_{meta} = 1.89 \times 10^{-12}$, OR $= 0.43$ for rs3077 with no evidence of heterogeneity ($P_{het} = 0.75$). Although the reason why heterogeneity was observed in replication-1 is unclear, one possible reason is the clinical heterogeneity due to different kits being used for antibody testing. The associations of *HLA-DPA1/−DPB1* with CHB and HBV clearance showed the same level of significance in the comparison of HBV patients with HBV resolved individuals (OR $= 0.43$ for rs3077 and 0.49 for rs9277542) as the one with healthy controls (OR $= 0.46$ for rs3077 and 0.50 for rs9277542), when the replication-1 was excluded in the analysis (Table 2 and Table 3). The results of meta-analysis across 6 independent studies including our study also showed the same or slightly weaker associations in the

Table 1. Number of study samples.

population		GWAS Japanese	Replication-1 Japanese	Replication-2 Korean
HBV carriers	Total	181	256	344
	IC	20	94	–
	CH	67	101	177
	LC	3	10	–
	HCC	91	51	167
Healthy controls		184	236	151
Resolved individuals		185	150	106

Abbreviation: IC, Inactive Carrier; CH, Chronic Hepatitis; LC, Liver Cirrhosis; HCC, Hepatocellular Carcinoma.
doi:10.1371/journal.pone.0039175.t001

Table 2. Results of replication study for protective effects against CHB.

dbSNP rsID	Chr	Position Buld 36.3	Nearest Gene	MAF[a] (allele)	Allele (1/2)	Stage (population)	HBV carriers 11	12	22	Healthy controls 11	12	22	HWEp	OR[b] 95% CI	P-value[c]	P_het[d]
rs3077	6	33141000	HLA-DPA1	0.44	T/C	GWAS	13	51	117	28	88	67	0.919	0.42	1.14×10^{-7}	
				(T)		(Japanese)	(7.2)	(28.2)	(64.6)	(15.3)	(48.1)	(36.6)		(0.30–0.58)		
						Replication-1	26	95	134	46	125	65	0.309	0.48	2.70×10^{-8}	
						(Japanese)	(10.2)	(37.3)	(52.5)	(19.5)	(53.0)	(27.5)		(0.37–0.62)		
						Replication-2	23	81	111	31	74	40	0.767	0.47	2.08×10^{-6}	
						(Korean)	(10.7)	(37.7)	(51.6)	(21.4)	(51.0)	(27.6)		(0.35–0.65)		
						Meta-analysis[e]								0.46	4.40×10^{-19}	0.80
														(0.39–0.54)		
rs9277542	6	33163225	HLA-DPB1	0.45	T/C	GWAS	18	53	110	29	102	52	0.073	0.42	5.32×10^{-8}	
				(T)		(Japanese)	(9.9)	(29.3)	(60.8)	(15.8)	(55.7)	(28.4)		(0.31–0.58)		
						Replication-1	30	106	118	54	114	67	0.681	0.54	3.33×10^{-6}	
						(Japanese)	(11.8)	(41.7)	(46.5)	(23.0)	(48.5)	(28.5)		(0.42–0.70)		
						Replication-2	30	87	94	35	72	36	0.933	0.54	8.29×10^{-5}	
						(Korean)	(14.2)	(41.2)	(44.5)	(24.5)	(50.3)	(25.2)		(0.40–0.74)		
						Meta-analysis[e]								0.50	1.28×10^{-15}	0.40
														(0.43–0.60)		

[a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19).
[b]Odds ratio of minor allele from two-by-two allele frequency table.
[c]P value of Pearson's chi-square test for allelic model.
[d]Heterogeneity was tested using general variance-based method.
[e]Meta-analysis was tested using the random effects model.
doi:10.1371/journal.pone.0039175.t002

comparison of HBV patients with HBV resolved individuals (OR = 0.56 for rs3077, 0.66 for rs9277535 and 0.55 for rs9277542) than in the one with healthy controls (OR = 0.55 for rs3077, 0.61 for rs9277535 and 0.51 for rs9277542), which was the opposite result as we expected (Table S2 and Table S4). These results may suggest that other unknown immune system(s) exist to eliminate the HBV in the HBV resolved individuals.

Among the HLA class II loci (HLA-DPA1, HLA-DPB1 and HLA-DQB2), which were associated with CHB and HBV clearance, a weak linkage disequilibrium ($r^2<0.1$) was observed between HLA-DQB2 locus and HLA-DPA1/−DPB1 loci in Japanese and Korean populations (Figure S2). We also found that similar linkage disequilibrium blocks (r^2) were observed among three subgroups (HBV carriers, HBV resolved individuals and Healthy controls). Moreover, logistic regression analysis of HLA-DP (rs3077 and rs92775542) with use of HLA-DQ (rs9276431 and rs768538) as covariates showed that the same level of significant associations of HLA-DP with CHB and HBV clearance as shown in the single-point association analysis, while no associations of HLA-DQ with $P_{log} >0.05$ were detected both in Japanese and in Korean (Table S5). These results show that HLA-DP is the main genetic factor for susceptibility to CHB and HBV clearance, and the associations of HLA-DQB2 would result from linkage disequilibrium of HLA-DPA1/−DPB1.

In this study, we confirmed the significant associations between HLA-DPA1 and HLA-DPB1, and protective effects against CHB and HBV clearance in Japanese and Korean individuals. These results suggest that the associations between the HLA-DP locus, CHB and HBV clearance are widely replicated in East Asian populations, including Chinese, Thai, Japanese and Korean individuals; however, there have been no similar GWAS performed in Caucasian and African populations. Moreover,

there were no significant SNPs associated with HCC development in this study, thus suggesting that it is necessary to increase the sample size. To clarify the pathogenesis of CHB or the mechanisms of HBV clearance, further studies are necessary, including a functional study of the HLA-DP molecule, identification of novel host genetic factors other than HLA-DP, and variation analysis of HBV.

Materials and Methods

Ethics Statement

All study protocols conform to the relevant ethical guidelines, as reflected in the a priori approval by the ethics committees of all participating universities and hospitals. The written informed consent was obtained from each patient who participated in this study and all samples were anonymized.

Genomic DNA Samples and Clinical Data

All of the 1,793 Japanese and Korean samples, including individuals with CHB, healthy controls and HBV-resolved individuals (HBsAg-negative and anti-HBc-positive), were collected at 20 multi-center hospitals (liver units with hepatologists) throughout Japan and Korea. The 19 hospitals in Japan were grouped into the following 8 areas: Hokkaido area (Hokkaido University Hospital, Teine Keijinkai Hospital), Tohoku area (Iwate Medical University Hospital), Kanto area (Musashino Red Cross Hospital, Saitama Medical University, Kitasato University Hospital, University of Tokyo), Koshin area (Shinshu University Hospital, Kanazawa University Hospital), Tokai area (Nagoya City University Hospital, Nagoya Daini Red Cross Hospital), Kinki area (Kyoto Prefectural University of Medicine Hospital, National Hospital Organization Osaka National Hospital, Osaka

Table 3. Results of replication study for clearance of hepatitis B virus.

	Position			MAF[a]	Allele	Stage	HBV carriers			Resolved individuals			OR[b]		
dbSNP rsID	Chr	Buld 36.3	Nearest Gene	(allele)	(1/2)	(population)	11	12	22	11	12	22	95% CI	P-value[c]	P_{het}[d]
rs3077	6	33141000	HLA-DPA1	0.44	T/C	GWAS	13	51	117	29	82	74	0.44	9.24×10^{-7}	
				(T)		(Japanese)	(7.2)	(28.2)	(64.6)	(15.7)	(44.3)	(40.0)	(0.32–0.61)		
						Replication-1	26	95	134	20	64	60	0.72	3.32×10^{-2}	
						(Japanese)	(10.2)	(37.3)	(52.5)	(13.9)	(44.4)	(41.7)	(0.53–0.97)		
						Replication-2	23	81	111	29	48	28	0.41	2.35×10^{-7}	
						(Korean)	(10.7)	(37.7)	(51.6)	(27.6)	(45.7)	(26.7)	(0.29–0.58)		
						Meta-analysis[e]							0.51	1.56×10^{-4}	0.03
													(0.36–0.72)		
						Meta-analysis[e]							0.43	1.89×10^{-12}	0.75
						(GWAS+replication-2)							(0.34–0.54)		
rs9277542	6	33163225	HLA-DPB1	0.45	T/C	GWAS	18	53	110	28	88	69	0.51	3.15×10^{-5}	
				(T)		(Japanese)	(9.9)	(29.3)	(60.8)	(15.1)	(47.6)	(37.3)	(0.37–0.70)		
						Replication-1	30	106	118	28	62	52	0.68	1.25×10^{-2}	
						(Japanese)	(11.8)	(41.7)	(46.5)	(19.7)	(43.7)	(36.6)	(0.51–0.92)		
						Replication-2	30	87	94	30	53	22	0.46	4.97×10^{-6}	
						(Korean)	(14.2)	(41.2)	(44.5)	(28.6)	(50.5)	(21.0)	(0.33–0.64)		
						Meta-analysis[e]							0.55	5.91×10^{-7}	0.19
													(0.43–0.69)		
						Meta-analysis[e]							0.49	9.69×10^{-10}	0.65
						(GWAS+replication-2)							(0.39–0.61)		

[a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19).
[b]Odds ratio of minor allele from two-by-two allele frequency table.
[c]P value of Pearson's chi-square test for allelic model.
[d]Heterogeneity was tested using general variance-based method.
[e]Meta-analysis was tested using the random effects model.
doi:10.1371/journal.pone.0039175.t003

City University), Chugoku/Shikoku area (Tottori University Hospital, Ehime University Hospital, Yamaguchi University Hospital, Kawasaki Medical College Hospital) and Kyushu area (Kurume University Hospital). Korean samples were collected at Yonsei University College of Medicine.

HBV status was measured based on serological results for HBsAg and anti-HBc with a fully automated chemiluminescent enzyme immunoassay system (Abbott ARCHITECT; Abbott Japan, Tokyo, Japan, or LUMIPULSE f or G1200; Fujirebio, Inc., Tokyo, Japan). For clinical staging, inactive carrier (IC) state was defined by the presence of HBsAg with normal ALT levels over 1 year (examined at least four times at 3-month intervals) and without evidence of portal hypertension. Chronic hepatitis (CH) was defined by elevated ALT levels (>1.5 times the upper limit of normal [35 IU/L]) persisting over 6 months (at least by 3 bimonthly tests). Liver cirrhosis (LC) was diagnosed principally by ultrasonography (coarse liver architecture, nodular liver surface, blunt liver edges and hypersplenism), platelet counts <100,000/cm³, or a combination thereof. Histological confirmation by fine-needle biopsy of the liver was performed as required. Hepatocellular carcinoma (HCC) was diagnosed by ultrasonography, computerized tomography, magnetic resonance imaging, angiography, tumor biopsy or a combination thereof.

The Japanese control samples from HBV-resolved subjects (HBsAg-negative and anti-HBc-positive) at Nagoya City University-affiliated healthcare center were used by comprehensive agreement (anonymization in an unlinkable manner) in this study. Some of the unrelated Japanese healthy controls were obtained from the Japan Health Science Research Resources Bank (Osaka, Japan). One microgram of purified genomic DNA was dissolved in 100 μl of TE buffer (pH 8.0) (Wako, Osaka, Japan), followed by storage at −20°C until use.

SNP Genotyping and Data Cleaning

For GWAS, we genotyped a total of 550 individuals, including 181 Japanese HBV carriers, 184 Japanese healthy controls and 185 spontaneously HBV-resolved Japanese individuals (HBsAg-negative and anti-HBc-positive), using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Inc., Santa Clara, CA), in accordance with the manufacturer's instructions. The average QC call rate for 550 samples reached 98.47% (95.00–99.92%), which had an average sample call rate of 98.91% (93.55–99.74%) by determining the genotype calls of over 900 K SNPs using the Genotyping Console v4.1 software (with Birdseed v1 algorithm) provided by the manufacturer [19]. We then applied the following thresholds for SNP quality control in data cleaning: SNP call rate ≥95% and MAF ≥1% for three groups (HBV carriers, healthy controls and HBV-resolved individuals), and HWE P-value ≥0.001 for healthy controls [20]. Here, SNP call rate is defined for each SNP as the number of successfully genotyped samples divided by the number of total samples genotyped. A total of 597,789 SNPs and 590,278 SNPs on autosomal chromosomes

passed the quality control filters in the genome-wide association analysis using HBV carriers and healthy controls, and using HBV carriers and HBV-resolved individuals, respectively (Figure 1). All cluster plots for the SNPs showing $P<0.0001$ on association analyses in the allele frequency model were confirmed by visual inspection, and SNPs with ambiguous cluster plots were excluded.

In the following replication stage, we selected a set of 32 SNPs with $P<0.0001$ in the GWAS using HBV carriers and HBV-resolved individuals. SNP genotyping in two independent sets of 256 Japanese HBV carriers, 236 Japanese healthy controls and 150 Japanese HBV-resolved individuals (Table 1, replication-1), and 344 Korean HBV carriers, 151 Korean healthy controls and 106 Korean HBV-resolved individuals (Table 1, replication-2) was completed for the selected 32 SNPs using the DigiTag2 assay [21,22] and custom TaqMan SNP Genotyping Assays (Applied Biosystems, Foster City, CA) on the LightCycler 480 Real-Time PCR System (Roche, Mannheim, Germany).

Statistical Analysis

The observed associations between SNPs and the protective effects on chronic hepatitis B or clearance of hepatitis virus B were assessed by chi-squared test with a two-by-two contingency table in allele frequency model. SNPs on chromosome X were removed because gender was not matched among HBV carriers, healthy controls and HBV-resolved individuals. A total of 597,789 SNPs and 590,278 SNPs passed the quality control filters in the GWAS stage; therefore, significance levels after Bonferroni correction for multiple testing were $P=8.36\times10^{-8}$ (0.05/597,789) and $P=8.47\times10^{-8}$ (0.05/590,278), respectively. For the replication study, 29 of 32 SNPs were successfully genotyped; therefore, we applied $P=0.0017$ (0.05/29) as a significance level, and none of the 29 markers genotyped in the replication stage showed deviations from the Hardy-Weinberg equilibrium in healthy controls ($P>0.01$).

The genetic inflation factor λ was estimated by applying the Cochrane-Armitage test on all SNPs and was found to be 1.056 and 1.030 in the GWAS using HBV carriers and healthy controls, and using HBV carriers and HBV-resolved individuals, respectively (Figure S3). These results suggest that the population substructure should not have any substantial effect on statistical analysis. In addition, the principal component analysis in a total of 550 individuals in the GWAS stage together with the HapMap samples also revealed that the effect of population stratification was negligible (Figure S4).

Based on the genotype data of a total of 1,793 samples including 1,192 Japanese samples and 601 Korean samples in both GWAS and replication stages, haplotype blocks were estimated using the Gabriel's algorithm in the Haploview software (v4.2) (Figure S2). In the logistic regression analysis, two SNPs (rs9276431 and rs7768538) within the HLA-DQ locus were individually involved as a covariate (Table S5). Statistical analyses were performed using the SNP & Variation Suite 7 software (Golden Helix, MT, USA).

Supporting Information

Figure S1 GWAS using samples from HBV carriers with LC or HCC, and HBV carriers without LC and HCC. P values were calculated using chi-squared test for allele frequencies. (PPTX)

Figure S2 Estimation of linkage disequilibrium blocks in HBV patients, HBV resolved individuals and healthy controls in Japanese and Korean. The LD blocks (r^2) were analyzed using the Gabriel's algorithm. (PPTX)

Figure S3 Quantile-quantile plot for test statistics (allele-based chi-squared tests) for GWAS results. Dots represent P values of each SNP that passed the quality control filters. Inflation factor λ was estimated to be: a) 1.056 in the analysis with HBV carriers and healthy controls; and b) 1.030 with HBV carriers and HBV-resolved individuals. (PPTX)

Figure S4 Principal component analysis on a total of 550 individuals in GWAS, together with HapMap samples (CEU, YRI and JPT). (PPTX)

Table S1 Results for 29 SNPs selected in replication study using samples of HBV carriers and healthy controls. [a]P values by chi-squared test for allelic model. [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]Meta-analysis was tested using additive, two-tailed CMH fixed-effects model. (XLSX)

Table S2 Results of meta-analysis for protective effects against persistent HB infection across 6 independent studies, including this study. [a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19). [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]P value of Pearson's chi-squared test for allele model. [d]Heterogeneity was tested using general variance-based method. [e]Meta-analysis was tested using the random effects model. (XLSX)

Table S3 Results for 29 SNPs selected in replication study using samples from HBV carriers and HBV-resolved individuals. [a]P values by chi-squared test for allelic model. [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]Meta-analysis was tested using additive, two-tailed CMH fixed-effects model. (XLSX)

Table S4 Results of meta-analysis for clearance of HBV across 6 independent studies, including this study. [a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19). [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]P value of Pearson's chi-squared test for allele model. [d]Heterogeneity was tested using general variance-based method. [e]Meta-analysis was tested using the random effects model. (XLSX)

Table S5 Logistic regression analysis of HLA-DP (rs3077 and rs9277542) and HLA-DQ (rs9276431 and rs7768538) with susceptibility to CHB and HBV clearance using the HLA-DQ genotypes individually as a covariate. (XLSX)

Author Contributions

Conceived and designed the experiments: NN HS YT. Performed the experiments: HS Y. Mawatari M. Sageshima YO. Analyzed the data: NN MK AK. Contributed reagents/materials/analysis tools: KM M. Sugiyama SHA JYP SH JHK KS M. Kurosaki YA SM MW ET MH SK EO YI EM AT Y. Murawaki YH IS M. Korenaga KH TI NI KHH YT MM. Wrote the paper: NN M. Kawashima YT KT MM.

References

1. Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO (2002) Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol 83: 2059–2073.
2. Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok ASF (2007) Management of hepatitis B: Summary of a clinical research workshop. Hepatology 45: 1056–1075.
3. Yokosuka O, Kurosaki M, Imazeki F, Arase Y, Tanaka Y, et al. (2011) Management of hepatitis B: Consensus of the Japan society of Hepatology 2009. Hepatol Res 41: 1–21.
4. Tada H, Uga N, Fuse Y, Shimizu M, Nemoto Y, et al. (1992) Prevention of perinatal transmission of hepatitis B virus carrier state. Acta Paediatr Jpn 34: 656–659.
5. Stevens CE, Toy PT, Taylor PE, Lee T, Yip HY (1992) Prospects for control of hepatitis B virus infection: implications of childhood vaccination and long-term protection. Pediatrics 90: 170–173.
6. Szmuness W (1979) Large-scale efficacy trials of hepatitis B vaccines in the USA: baseline data and protocols. J Med Virol 4: 327–340.
7. Kwon H, Lok AS (2011) Hepatitis B therapy. Nat Rev Gastroenterol Hepatol 8: 275–284.
8. Kobayashi M, Ikeda K, Arase Y, Suzuki F, Akuta N, et al. (2008) Change of Hepatitis B virus genotypes in acute and chronic infections in Japan. J Med Virol 80: 1880–1884.
9. Yano K, Tamada Y, Yatsuhashi H, Komori A, Abiru S, et al. (2010) Dynamic epidemiology of acute viral hepatitis in Japan. Intervirology 53: 70–75.
10. Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, et al. (2009) A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. Nat Genet 41: 591–595.
11. Mbarek H, Ochi H, Urabe Y, Kumar V, Kubo M, et al. (2011) A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. Hum Mol Genet 20: 3884–3892.
12. Li J, Yang D, He Y, Wang M, Wen Z, et al. (2011) Associations of HLA-DP variants with hepatitis B virus infection in southern and northern Han Chinese populations: a multicenter case-control study. PLoS ONE 6: e24221.
13. Guo X, Zhang Y, Li J, Ma J, Wei Z, et al. (2011) Strong influence of human leukocyte antigen (HLA)-DP gene variants on development of persistent chronic hepatitis B virus carriers in the Han Chinese population. Hepatol 53: 422–428.
14. Thursz MR, Kwiatkowski D, Allsopp CEM, Greenwood BM, Thomas HC, et al. (1995) Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N Engl J Med 332: 1065–1069.
15. Godkin A, Davenport M, Hill AVS (2005) Molecular analysis of HLA class II associations with hepatitis B virus clearance and vaccine nonresponsiveness. Hepatology 41: 1383–1390.
16. An P, Winkler C, Guan L, O'Brien SJ, Zeng Z, et al. (2011) A common HLA-DPA1 variant is a major determinant of hepatitis B virus clearance in Han Chinese. J Infect Dis 203: 943–947.
17. Wang L, Wu X-P, Zhang W, Zhu D-H, Wang Y, et al. (2011) Evaluation of genetic susceptibility loci for chronic hepatitis B in Chinese: two independent case-control study.
18. Hu L, Zhai X, Liu J, Chu M, Pan S, et al. (2011) Genetic variants in HLA-DP/DQ influence both hepatitis B virus clearance and Hepatocellular carcinoma development. Hepatology (in press).
19. Nishida N, Koike A, Tajima A, Ogasawara Y, Ishibashi Y, et al. (2008) Evaluating the performance of Affymetrix SNP Array 6.0 platform. BMC Genomics 9: 431.
20. Miyagawa T, Nishida N, Ohashi J, Kimura R, Fujimoto A, et al. (2008) Appropriate data cleaning methods for genome-wide association study. J Hum Genet 53: 886–893.
21. Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 364: 78–85.
22. Nishida N, Mawatari Y, Sageshima M, Tokunaga K (2012) Highly parallel and short-acting amplification with locus-specific primers to detect single nucleotide polymorphisms by the DigiTag2 assay. PLoS ONE 7: e29967.